

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
12. Mai 2005 (12.05.2005)

PCT

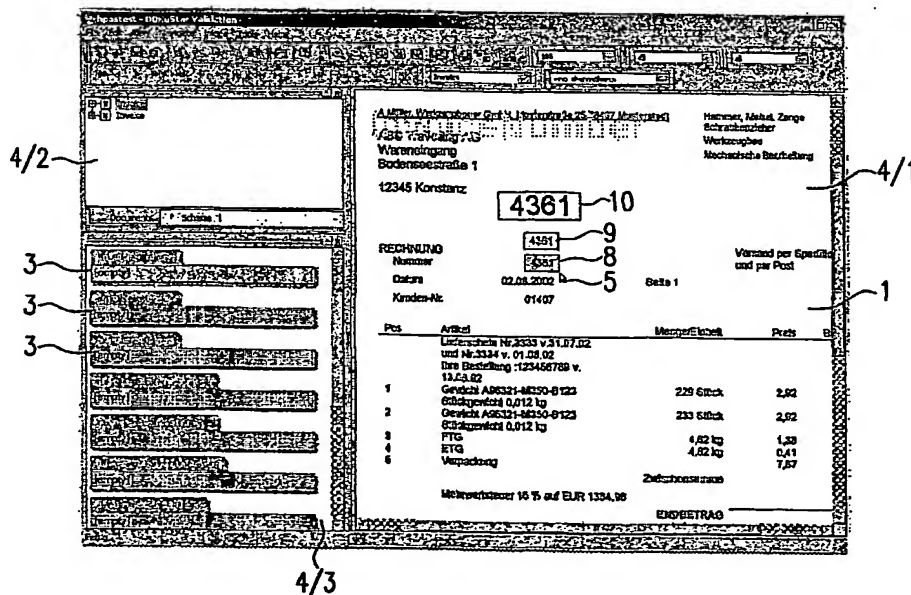
(10) Internationale Veröffentlichungsnummer
WO 2005/043452 A1

- (51) Internationale Patentklassifikation⁷: **G06K 9/20**
- (21) Internationales Aktenzeichen: **PCT/EP2004/009539**
- (22) Internationales Anmeldedatum:
26. August 2004 (26.08.2004)
- (25) Einreichungssprache: **Deutsch**
- (26) Veröffentlichungssprache: **Deutsch**
- (30) Angaben zur Priorität:
103 45 526.4 30. September 2003 (30.09.2003) **DE**
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **OCÉ DOCUMENT TECHNOLOGIES GMBH [DE/DE];** Max-Stromeyer-Strasse 116, 78467 Konstanz (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): **SCHIEHLEN, Matthias [DE/DE];** Kreuzackerweg 4, 78465 Konstanz (DE).
- (74) Anwälte: **SCHAUMBURG, Karl-Heinz usw.;** Postfach 86 07 48, 81634 München (DE).
- (81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): **AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.**

[Fortsetzung auf der nächsten Seite]

(54) Title: **METHOD AND SYSTEM FOR ACQUIRING DATA FROM MACHINE-READABLE DOCUMENTS**

(54) Bezeichnung: **VERFAHREN UND SYSTEM ZUM ERFASSEN VON DATEN AUS MASCHINELL LESBAREN DOKUMENTEN**



(57) Abstract: The invention relates to a method for acquiring data from machine-readable documents, whereby said data are associated with a database. According to the invention, string sections that are located close to a pointer moved by the user are automatically selected, highlighted and their content is proposed for transfer to the database. According to an embodiment of the inventive method, the content of a table can be read out in a fully automatic manner once the table entries in a first line of the table have been read out according to the aforementioned method.

[Fortsetzung auf der nächsten Seite]

WO 2005/043452 A1



(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— mit internationalem Recherchenbericht

— vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) **Zusammenfassung:** Die Erfindung betrifft ein Verfahren zum Erfassen von Daten aus maschinell lesbaren Dokumenten, wobei die Daten einer Datenbank zugeordnet werden. Mit der Erfindung werden Stringabschnitte, die sich in der Nähe eines vom Benutzer bewegbaren Zeigers befinden, automatisch ausgewählt, markiert und deren Inhalt zur Übernahme in eine Datenbank vorgeschlagen. Nach einer Weiterbildung des erfindungsgemässen Verfahrens kann der Inhalt einer Tabelle vollautomatisch ausgelesen werden, wenn die Tabelleneinträge in einer ersten Tabellenzeile einmal gemäss obigen Verfahren ausgelesen worden sind.